R E P O R T R E S U M E S

ED 010 355 48

LANGUAGES OF THE WORLD--IBERO-CAUCASION AND PIDGIN-CREOLE FASCICLE ONE.

BY- VOEGELIN, C. F. VOEGELIN, FLORENCE M.
INDIANA UNIV., BLOOMINGTON

REPORT NUMBER NDEA-VI-63-8 PUB DATE NOV 64
CONTRACT OEC-SAE-9488
EDRS PRICE MF-$0.18 HC-$3.08 77P. ANTHROPOLOGICAL LINGUISTICS, 6(8)/1-71, NOV. 1964

DESCRIPTORS- *PIDGIN, *LANGUAGES, *CAUCASIAN LANGUAGES, BLOOMINGTON, INDIANA, ARCHIVES OF LANGUAGES OF THE WORLD

LANGUAGES OF CAUCASIAN AREAS IN THE SOVIET UNION AND THE LINGUA FRANCAS OF THE MEDITERRANEAN REGIONS, THE ORIENT, AND THE WEST INDIES ARE LISTED AND DESCRIBED. (THIS REPORT IS PART OF A SERIES, ED 010 350 TO ED 010 367.) (JK)

ED010355

63-8

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

Anthropological Linguistics

Volume 6 Number 8

November 1964

Title

LANGUAGES OF THE WORLD:

IBERO-CAUCASIAN AND PIDGIN-CREOLE FASCICLE ONE

A Publication of the
ARCHIVES OF LANGUAGES OF THE WORLD
Anthropology Department
Indiana University

ANTHROPOLOGICAL LINGUISTICS is designed primarily, but not exclusively, for the immediate publication of data-oriented papers for which attestation is available in the form of tape recordings on deposit in the Archives of Languages of the World. This does not imply that contributors will be restricted to scholars working in the Archives at Indiana University; in fact, one motivation for the publication of ANTHROPOLOGICAL LINGUISTICS is to increase the usefulness of the Archives to scholars elsewhere by making publishable data and conclusions and their tape recorded attestation more widely available. (Recorded attestation of papers from scholars elsewhere will be copied by the Archives and the original recordings returned to the collector; others may then work with the tape copy either in the Archives or elsewhere by having a copy sent to them.) In addition to heavily exemplified papers in the form of preliminary or final statements on restricted problems in phonemics, morphophonemics, morphemics, syntax and comparative grammar, ANTHROPOLOGICAL LINGUISTICS will include lexical lists and analyzed texts (especially in the otherwise hard-to-publish range of 20 to 100 pages) and theoretical and methodological papers (especially in the form of papers from symposia).

Each volume of ANTHROPOLOGICAL LINGUISTICS consists of nine numbers to be issued during the months of January, February, March, April, May, June, October, November and December. Subscriptions ($3.50 a year) and papers for publication should be sent to the editor, Dr. Florence M. Voegelin, Anthropology Department, Indiana University, Bloomington, Indiana.

Entered as second class matter at the postoffice
at Bloomington, Indiana

LANGUAGES OF THE WORLD: IBERO-CAUCASIAN and PIDGIN-CREOLE FASCICLE ONE

C. F. and F. M. Voegelin

Indiana University

0. Scope note

1.0. Caucasian languages

1.1. Northwest Caucasian

1.2. Northeast Caucasian

1.3. South Caucasian

1.4. Caucasian structures

2. Modern Basque and extinct languages of the Mediterranean and the Most Ancient East

3. Pidgin-Creoles

The research reported herein was performed pursuant to a contract with the United States Office of Education, Department of Health, Education, and Welfare.

0. Iberian is used in at least two senses in linguistics. Most commonly, it refers to languages of the Iberian peninsula of southeast Europe (including Basque). But it also refers to the languages of southeast Asia--those in the Caucasus--which were once known as 'Japhetic languages'. But that was before Marr was discredited in Russia. In Marr's heyday, when active Russian linguists were followers of Marr, 'Japhetic languages' was a term either equivalent to 'Caucasian' or identifying Caucasian as a point of departure for far-flung and fantastic affiliations. Now 'Ibero-Caucasian' is sometimes used instead to refer to Georgian and other south Caucasian or Kartvelian languages ('Iberian') as distinguished from 'Caucasian' (Northwest and Northeast Caucasian).

Basque, now spoken in Spain and France, is often supposed to be distantly related to modern Caucasian languages, on the one hand, and to languages formerly spoken in the Most Ancient East on the other hand (2, below).

The connection between Caucasian (or languages of the Most Ancient East) and modern Pidgin-Creoles is non-existent, genetically speaking. There is a common problem that is shared by these two groups of languages, nevertheless. That problem has to do with a non-general tendency among natural languages, notably but not exclusively among Pidgin-Creoles.

The general tendency may be stated hypothetically. Suppose language A is in contact with a language of fewer speakers and less prestige, B; in consequence of a generation of such contact, B may borrow so much from A, that the next generation of speakers may be said to speak language A (relatively unchanged) and language B′(i.e. the former B, but now flooded with borrowing from A);

and the next generation, speaking B′, no longer speaks plain B.

In contrast to this general tendency, the non-general tendency of Pidgin-Creoles also begins with the contact of language A and language B; in consequence of this contact, some B speakers may adopt much of the vocabulary and some structural features of A, but retain the central structure and pronunciation of their native B if--and only if--speakers of the major language A imitate this new creation of a mixed language (C). Other B speakers may not even attempt to speak C, but may instead remain at home, as it were, speaking pure B. The only safe criterion of a mixed language (Pidgin-Creole) is that the unmixed languages A and B continue to be spoken beside the mixed language, C (the Pidgin-Creole). Thus, contacts between English (A) and Hawaiian (B) brought about an ephemeral Pidgin language (C) in which Christmas of language A is pronounced Kalikimaka in language C. When C (the Pidgin) was spoken, A (English) and B (Hawaiian) were also spoken; indeed A and B continue while C has levelled with A.

In most parts of the world, the general tendency is operative: contact between A (e.g. English before 1066) and B (e.g. Norman French) results in a reshaped A′ beside unchanged B, rather than the creation of a new language C beside the continuation of languages A and B. Cultural barriers or geographic discontinuity (e.g. the English Channel) inhibits leveling and favors this general tendency kind of continuing differentiation and continuation in two languages rather than the non-general convergence in a third language implied in pidgin-creoles. In the usual unilateral bilingual situation, the result is unchanged B beside reshaped A′. But the result of bilateral bilingualism is reshaped A′ beside reshaped B′.

Though neither the modern languages of the Caucasus (spoken east of the Black Sea) nor the extinct languages of Europe or of the Most Ancient East (spoken south and west of the Black Sea) are supposed to be examples of pidgin-creoles, they may still contain (or have contained) instances of pairs of languages (as A and B) in a kind of contact in a post-neolithic culture in which employers or owners (speaking A) would imitate workers (speaking B) and thereby converge or mix languages A and B to form a new language, C. And if A and B were unrelated, how would language C be related to them? In the same language family? the same phylum? or neither?

CAUCASIAN LANGUAGES

1.0. Caucasian languages are spoken primarily in the Caucasus area of the Soviet Union by about five million people. The Caucasus area in which these languages are spoken stretches from the Black Sea to the Caspian Sea (a distance of some 300 miles) and is flanked by Indo-European languages spoken both in the country to the south (Iran), and also in Russia north of the Kuban River. The south to north distance (from Iran to the middle reaches of the Kuban) is about 600 miles.

The language situation is complex enough when only Caucasian languages are considered, since they are numerous and widely divergent; in addition they are spoken in an area where they are in contact with languages of four different language families wholly unrelated to Caucasian--and the languages of all five families are spoken in an area little larger than Java, for example. However, the Caucasian languages have the best claim to being the aboriginal languages of the Caucasus (but the Caucasus are sometimes--though dubiously--claimed as an Uralic homeland, preceding the Ural mountain homeland). In the historical period the number of different Caucasian languages spoken in the area is known to exceed those belonging to other families. Beside some thirty odd Caucasian languages in the Caucasus, there are about a dozen Indo-European languages (Russian, Ukrainian, Polish, Bulgarian, German, Greek, Armenian, Ossetic, Tat, Talysh, Kurdish, Persian), and several Turkic languages (Nogai, Kumyk,

Karachai, Balkar, Turkmen, Azerbaijani, Osmali, Karapapak); there are also three or four Finno-Ugric enclaves and Semitic enclaves in the Caucasus, as well as Kalmuck, a Mongol language.

The Caucasian languages are generally divided into three 'groups':

Northwest Caucasian (4 languages)

Northeast Caucasian (27 languages)

South Caucasian (3 languages).

The relationship between these three 'groups' is not at all clear. If Caucasian as a whole were a language family in the sense of being reconstructable in detail (comparable to the reconstructions of Indo-European or of Finno-Ugric, for example), then the three 'groups' would be counted as branches of the Caucasian family, much as Slavic, Germanic, and Italic are counted as branches of Indo-European. If, on the other hand, the three were so very remotely related to each other that only an occasional or sporadic reconstruction were possible, then each 'group' would be counted as a language family, with the three 'groups' related in a Caucasian phylum. It is questionable whether enough descriptive material on individual languages in the three 'groups' has been analyzed to permit a final answer to the alternative interpretation.

Aert H. Kuipers (Current Trends in Linguistics Volume 1, 1963, p. 315) makes the following statement concerning the internal relationships of Caucasian: "The existence of a genetic relationship between NW and NE Caucasian is probable; the relations of S Caucasian to this N group so far remain unclear, though many Soviet linguists take a genetic relationship for granted, or at least start from it

as a working hypothesis."

N. S. Trubetzkoy and others have proposed a number of reasonable etymologies in an attempt to relate all of the Caucasian languages. But in general, Caucasian comparative studies are still in a state of infancy, and as John Lotz pointx out in the preface to Peoples and Languages of the Caucasus (1955), "It must be borne in mind that the exact genetic relationships of the Caucasian languages . . . are obscure."

"The most urgent task in Caucasian linguistics [in the post-Marrist period] remained the recording of facts; the description of the grammatical structure and the collection of lexical material of languages and dialects so far not or insufficiently studied. Two reasons made this task particularly urgent. In the first place, due to the prestige of the literary languages and of Russian, combined with changes in culture and increased mobility of the population, more and more dialect-material threatened to be irretrievably lost. In the second place, this task had to be fulfilled before comparative work could be undertaken with any chance of success. The N Cauc. languages - especially the NW group - put great obstacles in the way of phonetic comparison. This is due to the fact that the roots of these language in general consist of a single consonant (or a harmonic complex); as a result, each morpheme tends to present a unique combination of features, so that it is difficult to find series of comparisons involving the same sound (in principle, this is possible only in the case of homonyms). . . . Furthermore, it is in general the non-literary dialects that are phonetically the most complex, i.e. these dialects preserve distinctions lost in the literary dialects. The non-literary dialects

are therefore of the utmost importance for the reconstruction of the proto-forms of the various dialect-groups; such a reconstruction is necessary before one can fruitfully start comparing one dialect-group with the next. It is clear, therefore, that the description of dialects with a view to their importance for inter-dialect comparison is an indispensable prerequisite for the further development of comparative Cauc. linguistics" (Kuipers, p. 322).

The proportion of preliterate to literary languages is about 3 to 1. The following eleven Caucasian languages have official status as literary languages in the Soviet Union:

Georgian (2,700,000 speakers)

Chechen (419,000)

Lezghian (223,000)
Avar (270,000)

Kabardian (204,000)

Dargwa (158,000)

West Circassian (155,000)

Ingush (106,000)

Abxazo, i. e., the Abzhui dialect of Abkhaz,(65,000)

Lak (64,000)

Tabasaran (35,000)

Abaza, i. e., the Tapanta dialect of Abkhaz,(20,000)

Only Georgian has an old literary tradition with an alphabet of its own (identified below); the others in the preceding list use the Cyrillic alphabet. What ever else may be meant by the term 'literary language', the term is

applied to a language which is written in an alphabet that is used in the schools where the language is spoken.

NORTHWEST CAUCASIAN

1.1. The three or four Northwest Caucasian languages are spoken by nearly a half million speakers located in widely scattered areas on both sides of the main chain of mountains from the Kuban River to Abkhazia to North Osetia. West Circassian is spoken not only in the Caucasus, but also by immigrant groups in Turkey, Syria, Jordan, Israel, and Iraq.

(1) Abxazo (Abkhaz)-Abaza (Abazin (as a whole; Abkhaz) is spoken by 85,000 people in two areas in the USSR: the Abkhaz ASSR south of the mountains along the coast of the Black Sea; and in fifteen villages located north of the mountains in the Cherkes AO (near the Kuban headwaters). The latter are known collectively as Abaza, which does not represent a linguistic group, since the Ashxar dialect is closely similar to the dialects south of the Caucasus ridge, while Tapanta is very divergent. In addition to the Abkhaz speakers in the USSR, 8,602 speakers were reported in the Turkish census of 1945. Many Abkhaz speakers in the Abkhaz ASSR also speak Georgian and all speakers of the Abazo Abkhaz dialects are said to be bilingual, speaking also East Circassian. The five major dialects are:

Bzyb

Abzhui

Samurzakan

Ashxar (Ashkarwa, also called Abaza along with Tapanta)

Tapanta (Abaza), the most divergent dialect.

(2) Ubyx, (Ubykh), formerly spoken north of Abxaz along the Black Sea

coast, is now spoken only in Turkey; it is said to be becoming extinct.

(3) Circassian (Adyghe, Adygei, Cherkes) is divided into two main dialect groups each represented by a separate literary language; in terms of degree of mutual intelligibility it may be that these dialect groups represent two separate, but closely related, languages rather than dialects of the same language. The total number of Circassian native speakers is around 275,000. The Russian census of 1959 treats the Circassians as three groups: Kàbardiàn (those in the Kabardian ASSR), 204,000 of whom 95.0 per cent report Kabardian as their native language; the Adygei (those in the Adygei AO), 80,000 of whom 96.8 per cent are reported as native speakers; and Cherkes, 30,000, of whom 89.7 per cent are reported as native speakers. The 1945 census of Turkey reported 66,691 Circassian speakers, and there are smaller groups of Circassians in Syria, Jordon, Israel and Iraq.

(3a) West Circassian (Lower Circassian, Kiakh, Kjax) is spoken in numerous small enclaves scattered along the Kuban River and its tributaries in the Adygei AO; it is also spoken southeast of Tuapse and on the Taman Peninsula.

West Circassian dialects in the Caucasus include:

Shapsug, including Xakuchi

Bezhedux (Bzhedukh)-Temirgoi (Chemgui)

Abadzex (Abadekh)

Natuxaj (Natukhai).

(3b) East Circassian (Upper Circassian, Kabardian) is spoken in the Kabardian ASSR, between the upper Kuban and Zelenchuk rivers in the

Cherkes AO, and in the eastern part of the Adygei AO; the main area is located immediately west of North Osetia. East Circassian dialects include:

Greater Kabardian

Baksan

Lesser Kabardian

Malka

Mozdok

Kuban

Cherkes

Beslenei (sometimes said to be a transitional dialect between East and West Circassian).

NORTHEAST CAUCASIAN

1.2. The Northeast Caucasian languages are generally divided into four major groups--subgroups, from the point of view of all Caucasian languages.

(1) Vejnax (3 languages); these are geographically northernmost.

(2) Avaro-Andi-Dido (12 languages); these are north central geographically.

(3) Lak-Dargwa (2 languages); these are south central.

(4) Lezghian or Samur (10 languages); these are geographically southernmost.

For the total of twenty seven Northeast Caucasian languages one may compute the total number of speakers of these languages--namely, one and a quarter million people living for the most part in the Daghestan ASSR, near the Caspian coast; and compare this total with the total number of speakers of Northwest languages, above (less than half a million) and with South Caucasian, below (three million).

VEJNAX SUBGROUP OF NORTHEAST CAUCASIAN

There are three languages in the Vejnax (Veinakh, Samurian, Kist), subgroup, two of which are literary languages.

(1) Chechen is spoken by 419,000 people. Once located in the now abolished Chechen-Ingush ASSR, a large number of Chechen (and Ingush) speakers now live in the Kazakh SSR; a few are found in the Georgian SSR. The following five Chechen dialects are distinguished (not including Ingush

(2, below) which is sometimes treated as a dialect of Chechen rather than as a separate language):

Greater Chechen

Itumkala

Cheberloj

Galanchog

Akka (Aux), a transitional dialect between the Chechen language (to which it belongs) and the following Ingush language; possibly a separate language.

(2) Ingush is spoken by 106,000 people who are located among the Chechen--(1), above.

(3) Bats is spoken by 2,500 people in the Georgian SSR near the source of the Andi Koissu River. This dialect is linguistically less close to Chechen and Ingush than the latter two are to each other. Georgian is used as a literary language by the Bats.

AVARO-ANDI-DIDO SUBGROUP OF NORTHEAST CAUCASIAN

The dozen languages in this subgroup are located in an area of the southern Sulak River, and its tributaries (Andi Koissu, Avar Koissu and Kazikumukh Koissu). Avar serves as the literary language of people speaking all the other languages in the Avaro-Andi-Dido subgroup.

(1) Avar is spoken by 270,000 people who live both in the Daghestan ASSR and also, to the south, in the Azerbaidzhan SSR. The Avar dialects and sub-dialects are:

Northern, including: Salatav and Xunzax (Khunzakh)

Transitional, including:

Keleb

Bacadin

Untib

Shulanin

Kaxib

Southern, including:

Hid

Andalal-Gidatl

Karax (Karakh)

Ancux (Antsukh)

Batlux

Zakataly (Char)

The Andi sub-group comprises languages numbered (2), through (8), below. They are all spoken in Daghestan, ASSR.

(2) Andi proper is spoken by 8,000 people in dialects known as:

Munin

Rikvani

Kvanxidatl

Gagatl

Zilo.

(3) Botlix (Botlikh) and Godoberi are dialects of one language spoken by

an estimated 3,500 people.

(4) Karata is estimated to have 6,000 speakers, including those who speak the dialects Tokita and Anchix.

(5) Tindi is spoken by 5,000 people.

(6) Bagulal (Kvanada) is spoken by 5,500 people in two dialects:

Bagulal proper

Tlisi.

(7) Chamalal is spoken by 5,500 people in two dialect groups:

Gakvari, including Agvali, Tsumada, Tsumada-Urukh, Richaganik,

-Gadyri, including Gachitl,-Kvankhi

Gigatl.

(8) Axvax (Akhvakh) is spoken by an estimated 3,000 people including those who speak the Kaxib dialect.

Languages numbered (9) through (12) below constitute the Dido (Tsez, Cez) subgroup.

(9) Dido (Tsez, Cez) is spoken by an estimated 7,000 people in the Daghestan ASSR.

(10) Xvarshi (Khvarshi) is spoken in two dialects by an estimated 1,800 people in the Daghestan ASSR: Xvarshi and Inxokvari.

(11) Ginux (Ginukh) is spoken by some 200 people in the Daghestan ASSR.

(12) Bezhita (Kapucha)-Gunzib (Xunzal, Khunzal) is spoken by 1,600 people in the Daghestan ASSR.

LAK-DARGWA SUBGROUP OF NORTHEAST CAUCASIAN

Lak and Dargwa are the only two languages in this group (subgroup); these are spoken in an area of Daghestan ASSR which extends for forty miles from the southern Kazikumukh Koisu: to the south and to the east.

(1) The five Lak dialects are spoken by a total of 64,000 people:

Vicxin (Vitskhin)

Kumux (Kumukh)

Vixlin (Vikhlin)

Ashtikulin

Balxar-Calakan. (Balkhar-Tsalakan).

(2) The nine Dargwa dialects are spoken by a total of 158,000 people:

Cudaxar (Tsudakhar)

Uraxa-Axusha

Urkarax (Urakha-Akhusha, Akkhusha)

Kajtak (Xajdak, Kaitak)

Kubachi

Dejbuk

Xarbuk

Muirin

Sirxin.

LEZGHIAN (SAMUR) SUBGROUP OF NORTHEAST CAUCASIAN

The ten languages which follow represent the southernmost of the four subgroups of Northeast Caucasian; they are spoken not only in Southern Daghestan but also in adjacent parts of Azerbaidzhan SSR, and the Georgian SSR. The inclusion of the last two languages--those numbered (9) and (10), below--in this Lezghian (Samur) subgroup is still uncertainly attested.

(1) Lezghian dialects and subdialects are spoken by a total of 223,000 people in the Daghestan ASSR and the Azerbaidzhan SSR:

Kjuri (Kiuri)

Axty (Akhty)

Kuba

Gjunej

Garkin, including:

Anyx

Staly.

(2) Four Agul dialects are spoken by 7,000 people in the Daghestan ASSR:

Agul proper

Keren

Koshan

Gekxun.

(3) Rutul is spoken by perhaps 10,000 people both in the Daghestan ASSR and in the Azerbaidzhan SSR. Lezgian is used as a literary language by the Rutul. The dialects of Rutul include:

Shina

Ixreko-Muxrek

Borch.

(4) Tsaxur (Tsakhur, Caxur) is spoken by perhaps 11,000 people in the Daghestan ASSR and the Azerbaidzhan SSR. The Tsaxur use Avar as a literary language in the Daghestan ASSR and Azeri in the Azerbaidzhan SSR. Tsaxur dialects include:

Kirmitso-Lek

Mikik

Mishlesh.

(5) Southern and Northern Tabasaran dialects are spoken by 35,000 people in the Daghestan ASSR. Northern Tabasaran is also called Khanag.

(6) Budux (Budukh) is spoken by 2,000 people in the Azerbaidzhan SSR (1926 census).

(7) Kryts (Dzhek, Kryc, Kryz) is spoken by an estimated 2,600 people in the Azerbaidzhan SSR. Azeri is used as a literary language by the Kryts, the Budux, the Udi and the Xinalung. Kryts dialects are:

Kryts

Dzhek

Xaput (Khaput).

(8) Udi is spoken by an estimated 2,000 people in the Azerbaidzhan SSR and the Georgian SSR. Udi dialects include:

Vartashen

Nidzh (Nizh).

(9) Xinalug (Khinalug) was spoken by 1,500 people in the Azerbaidzhan SSR according to the 1926 census; a more recent estimate gives only 100 as the number of speakers.

(10) Archi is spoken by 859 people in the Daghestan ASSR (1926 census).

SOUTH CAUCASIAN (KARTVELIAN)

1.3. The most famous South Caucasian language is Georgian, which is the only Caucasian language with over a million speakers. Indeed, there are more nearly three million speakers of Georgian, which is the oldest written Caucasian language; Georgian is also used as the literary language of speakers of other South Caucasian languages. Evidence points to a 5th century introduction of the Greek alphabet by Christian missionaries who may have used or been influenced by an earlier Aramaic script. The 10th century shape of the Georgian alphabet was angular; the round modern shape of letters is called mkhedruli.

(1) Georgian is spoken not only in the Georgian SSR, but also in the Azerbaidzhan SSR, as well as in adjacent parts of Turkey and Iran--by 2,700,000 people, altogether. The dialects and subdialects of Georgian are:

Eastern, including:

Kartlian

Kaxetian (Kakhetian)

Ingilo

Tush

Xevsur (Khevsur)

Moxev (Mokhev)

Pshav

Mtiul

Ferejdan (in Iran)

Western, including:

Imeretian

Racha

Lexchxum (Lechkhum)

Gurian-Adzhar

Imerxev (in Turkey).

Georgian as a whole, however, is said to be dialectically quite homogeneous--more so than other dialectically differentiated Caucasian languages.

(2) Zan is spoken by an estimated 349,000 people in the northwestern part of the Georgian SSR north of the Rion River and in and around the cities of the Abkhaz ASSR. Zan is divided into two major dialect groups, Mingrelian (300,000 speakers) and Laz (2,000 speakers in USSR, 47,000 in Turkey). These are subdivided as follows:

Mingrelian:

Samurzakan-Zugdidi (Western)

Senaki (Eastern)

Laz (Chan):

Eastern, including:

Xopa (Hopa)

Chxala (Chkhala) .

Western, including:

Vice-Arxava (Vitse-Arkhava)

Atina.

(3) Svan is spoken by between 13,142 (1926 census) and 23,000 people (1944 estimate) in the mountainous regions of the northwestern Georgian SSR. Dialects include:

Upper Bal

Lower Bal

Lashx

Lentex.

CAUCASIAN STRUCTURES

1.4. The vowel systems of the Northwest Caucasian languages show (in terms of the number of phonemic contrasts) the simplest known systems in the world: one distinction on each of two tongue heights, / a/ (lower) and / ə/ (higher). Each of these two contrastive vowels has a large variety of non-contrastive allophones (lower front and back for / a/ and higher front and back for / ə/) predictable in terms of the surrounding consonants. The higher vowel / ə/ has often been interpreted as a junctural feature. In several of the Northwest languages there is a third vowel generally analyzed as long / a:/ (which appears to be roughly at the same tongue height as / a/).

Abxazo-Abaza (Tapanta dialect)	Kabardian, Ubyx
ə	ə
a	a
	a:

The vowel systems of the other Caucasian languages (Northeast and South) have at least five contrastive vowel qualities combining variously with components of length, nasalization and pharyngealization (these components are especially conspicuous in the Northeast languages).

The vowel systems may be simple, as indicated; the consonant systems compensate or overcompensate for paucity of distinctions: they go to the opposite extreme. Thus, the consonant systems (more accurately the stops, affricates and fricatives) of the Northwest Caucasian languages are among the most complex of all known languages of the world. According to Hans Vogt (Dictionnaire de la Langue Oubykh, Oslo, 1963) Ubyx distinguishes 80 consonants. According to W.S. Allen (Structure and System in the Abaza Verbal Complex, 1956), the Tapanta dialect of Abxazo-Abaza distinguishes 63 consonants. According to Aert H. Kuipers (Phoneme and Morpheme in Kabardian, 1960), the Kabardian dialect of East Circassian distinguishes 49 consonants.

The Ubyx system utilizes five series generating components: glottalization, voicing, pharyngealization, palatalization and labialization. Stripped of these components, this system is still far from simple, since it makes linear distinctions among seven fricatives and also among seven stops (including affricates among the stops):

p	ċ	t	c	č	k^y	q
f	ṡ	s		š	x	ẋ	h
m		n				
		l, ɬ, x^l				
		r				
w			y			

The linear distinctions would be eight, if /ɬ/ were counted as an additional fricative, and /ƛ'/ as an additional stop. Special letters, /c̈/ and /s̈/, are used for sounds which contrast with /c/ and /s/, the former two being tongue-tip up and fronted--comparable to t^{θ} and θ respectively--and the latter two being tongue-tip down; /ɬ/ is a voiced lateral fricative, /ɬ/voiceless, and /ƛ'/ affricated-glottalized.

When the series generating components already mentioned are combined with the plain consonants as abstracted above, no additional linear distinctions are made; but additional unit phoneme contrasts appear in the full Ubyx consonant system which is now given (with underlining to mark pharyngealization):

p	p̲	c̈	t	t^{w}	c	c^{w}	c̈	c̈^{y}	k^{y}	k^{w}	q	q^{y}	q^{w}	q̲	q̲^{w}
p'	p̲'	c̈'	t'	$t^{w\circ}$	c'	$c^{w\circ}$	c̈'	c̈$^{y\circ}$	$k^{y\circ}$	$k^{w\circ}$	q'	$q^{y\circ}$	$q^{w\circ}$	q̲'	q̲$^{w\circ}$
b	b̲	ż	d	d^{w}	ʒ	z^{w}	ʒ̈	ʒ̈^{y}	g^{y}	g^{w}					
f	f̲	s̈		s	s^{w}	s̈	s̈^{y}	s̈^{w}	x		ẋ	ẋ^{y}	ẋ^{w}	x̲	x̲^{w}	h
		ż		z	z^{w}	z̈	z̈^{y}	z̈^{w}	ɣ		ɣ̇	ɣ̇^{y}	ɣ̇^{w}	ɣ̲	ɣ̲^{w}	
m	m̲	n														
		l, ɾ, ƛ'														
		r														
w	w̲							y								

The five series generating components (mentioned above) introduce 58 phonemic units in addition to 22 simple consonants (counting /k^{y}/ as well as /k̈/ as simple consonants since there is no phonetically simpler [k] or [ƛ]). In distinctive feature analysis as developed by Roman Jakobson and others,

the Ubyx phonemic system (as presented above--that is to say, without breaking consonants like /c^w/ into a series consisting of /t/ plus /s/ plus /w/ plus */ʔ/, etc.) all twelve of the postulated features would be involved as contrastive.

One of the grounds for analyzing so many unit phonemes in Caucasian languages in general--and scholars agreed on this kind of analysis-- appears to be the great variety of consonant clusters that occur even with such an analysis. Such analysis also avoids excessive morphophonemic complexities.

Many Caucasian languages have a good number of pharyngeal and laryngeal stops and fricatives; a few also have glottalized fricatives, e.g. Kabardian (Kuipers, 1960) with both of these features:

p	t	c		k	k^w	q	q^w	ʔ	$ʔ^w$			
p'	t'	c'		k'	$k^{w'}$	q'	$q^{w'}$					
b	d	ʒ		g	g^w							
f	ṡ	s	š	x	x^w	ẋ	$ẋ^w$	h		ḥ	$ḥ^y$	$ḥ^w$
v	ż	z	ž	ɣ		ɣ̇	$ɣ̇^w$					
f'		s'										
	ɬ, ɬ', ɬ'											
m	n											
	r											

(Note: /ɬ/ is a voiced lateral fricative, /ɬ̣/ voiceless.)

The phonemic system of modern Georgian (after Deeters, Encyclopaedia Britannica 10, 1964) is among the simplest of all Caucasian languages:

p	t	c	č	k			i	u
p'	t'	c'	č'	k'			e	o
b	d	ʒ	ǯ	g			a	
		s	š	x	h			
v		z	ž	ɣ				
m		n						
		l						
		r						
			y					

The plain stop series is aspirated and the second series unaspirated and sometimes glottalized.

Proto-Daghestan (i.e. NE minus the Vejnax group) as reconstructed by E. Bokarev (Introduction to the Comparative-Historical Study of the Daghestan Languages [in Russian], Maxachkala, 1961, p. 58, cited by Kuipers) is as follows:

p	t	c	č	k	q	q̌
p'	t'	c'	č'	k'	q'	q̌'
b	d	ʒ	ǯ	g		ǧ
		s	š	x		x̣
		z	ž	ɣ		ɣ̣
m		n				
		l, ɬ, ɬ', λ				
		r				

Note: /ǫ̌ ǫ̌' ǧ/ are affricates.

Boxed consonants also occur geminated. The Proto-Daghestan vowel system is:

i	u
e	(o)
æ	a

probably combining with a component of pharyngealization as well as nasalization.

Proto-South Caucasian can be reconstructed with much greater certainty than Daghestanian. G. A. Klimov (Opyt rekonstrukcii fonemnogo sostava obščekartvel'skogo jazyka-osnovy, Izvan 1960, pp. 22-31, cited by Kuipers) has made the following reconstruction:

p	t	c	č	k	q	i	u
p'	t'	c'	č'	k'	q'	e	o
b	d	ʒ	ǯ	g	(ġ)	a	
		s	š	x			
		z	ž	ɣ			
m	n						
	l						
	r						
			y				

The dental and palatal affricates and fricatives cause some difficulties which have not yet been fully solved.

Northwest Caucasian phoneme correspondences have not yet been worked out in sufficient detail to warrant an attempt at rigorous reconstruction.

Our report focuses on the diversity of consonant systems among Caucasian languages; Kuipers (1963) is especially impressed by the diversity of vowel systems, and offers generalizations for the rest of Caucasian grammar (p. 318, 319):

"While the NW Cauc. languages have simple case-systems and a rich, polypersonal conjugation, NE Cauc. has extensive, serial case-systems and mostly does not express person in the verb. The S Cauc. languages occupy in both respects an intermediate position. In many of the NE Cauc. languages morphology and syntax are dominated by a division of nouns into classes (rational, subdivided into male and female, and one or more non-rational classes). The verb contains class-indicators rather than person-affixes, e.g. Avar sing. rat. masc. w, fem. j, irrat. b; plur. r, l (all classes): dun w-ač'una "I (man) come", dun j-ač'una "I (woman) come", ču b-ač'una "the horse comes", niž r-ač'una "we come", wac: roq: 'o-w w-ugo "brother at-home is", jac: roq: 'o-j j-igo "sister id.", t'ex: roq: 'o-b b-ugo "book in-the-house is", roq: 'o-r r-ugo "they are at home." Lexical elements may contain petrified class-prefixes, cf. wac: "brother", jac: "sister", rag "battle" (root -ağ-, cf. b-ağ-ize, etc. "to quarrel"), roq: ' "house."

"All the Cauc. languages have different cases for the subject of an intransitive and for the actor of a transitive verb (secondary modifications in S Cauc.); the trans. actor is expressed either by a special "ergative" case

or by a general "oblique" case, the intrans. subject and the goal of a trans. verb being expressed by a casus rectus (also called "nominative" or "absolutive"), so that from a Caucasian point of view the latter two categories are one and the same."

2.0. The known or partially known non-Indo-European and non-Semitic languages and language families (in Mediterranean Europe and in Anatolia of the Most Ancient East) include Basque (2.1, below), which is still spoken, and a score of extinct languages from Iberian (2.2) in the Spanish peninsula, to Etruscan in Italy (2.3), to languages formerly spoken on islands in the eastern Mediterranean (2.4 - 2.7), to Sumerian in south Mesopotamia (2.8) and Elamite in Persia (2.9), and to lesser known languages in Asia Minor (2.10 ff).

2.1. Basque is the only living language of Western Europe that does not belong to the Indo-European family.

Numerous attempts have been made to relate Basque to other languages, including the Hamitic languages--Cushitic and Berber--of North Africa; the Dravidian languages of India; the Algonquin languages of North America; and the extinct European languages, especially Ligurian and Iberian.

Such specialists in Basque research as René Lafon and C. C. Uhlenbeck, however, believe Basque to be direct descendants of the Aquitanians (who in the time of Julius Caesar inhabited the region between the Garonne River, the Pyrenees and the Atlantic) and the Vascones (who once occupied the greater part of the Spanish Navarre and adjacent areas). The self-designation of the Basques is Euskara, Eskuara or Üskara (according to the various dialects). This designation, it will be noted, is strikingly similar to the Aquitanian term, Ausci.

The only possibility of a significant broader relationship between Basque

and other living languages, according to Lafon, Uhlenbeck and others, involves Caucasian. If this is true, the Basque forebears may have come from the Caucasus four or five millenia ago.

Basque is now spoken in the western Pyrenees Mountains on both sides of the Franco-Spanish border. The Basque speakers in this part of Europe number 600,000 or 700,000 of whom about 90,000 live in France; unknown numbers of Basque have emigrated to the New World, largely in their capacity as shepherds.

There are eight Basque dialects, known mainly by French and/or Spanish provincial names:

Viscayan (most divergent as well as most widely spoken)

Guipúzcoan

Northern Upper Navarrese

Southern Upper Navarrese

Labourdin

Western Lower Navarrese

Eastern Lower Navarrese

Souletin.

The first four of the dialects listed above are spoken entirely in Spain; the last four are spoken mainly in France, but each also has subdialect representatives in Spain.

The vowel system of Basque is of the type 2(FB) over N--/i u/ over /e o/ over /a/--to which must be added a front rounded vowel /ü/ in the Souletin

dialect. The Basque consonant system (Guipúzcoan dialect), as described by Dorita Lochak (AL 2.3. 12-31, 1960) is as follows:

p	t	c	č	k
b	d			g
	s		š	x
m	n		n^y	
	l		l^y	
	ɾ			
	r			
w			y	

2.2. Iberia was the common Greek name for the Spanish peninsula. There is also a narrower meaning for Iberia--the people or country along the Ebro River (Iberus in Latin, ibay-erri the country of the river in Basque): Iberian was the language of the peninsula, in the broader sense; and in the narrower sense, Iberian was the language of the Ebro River people. But Iberian was only one of the several languages known in the peninsula; it was known to be spoken as far east as the Rhone River by the time of Strabo.

Today, fragments of the Iberian language are preserved in many coins (and few inscriptions). The coins are found in various parts of the peninsula (e.g. near Narbonne). The Iberian alphabet may represent an older script reshaped under Greek and Punic alphabetic influence. By a combination of Iberian and Basque derivations, K. W. von Humboldt accounts for many place names in the peninsula, and conjectures that the modern Basque (2.1,

above) are descendants of Iberian speakers (2.2) who may have once inhabited neighboring islands (Sardinia, Corsica, and even the British Isles), as well as the peninsula and adjacent France. This Iberian theory enjoyed greater popularity in physical anthropology than in historical linguistics.

2.3. In the diffusion of our type of alphabet which specifies vowels as well as consonants, Etruscan was an intermediary between the Greek alphabet as donor, and the later Latin alphabet. The alphabetic values of Etruscan writing are known, but the Etruscan morphemes (other than borrowings) cannot yet be read; in this sense, Etruscan remains an undeciphered language, even though the available material written in Etruscan far exceeds that of Mycenean Linear B, an earlier Greek dialect which was deciphered in 1952. Etruscan is known to have been spoken in Italy from the 8th century B.C. to the 4th century A.D., when it became extinct. The language is preserved in one long text (1500 words written on the wrapping of a mummy now in the Zagreb Museum), in a few short texts (one to three hundred words each), and in some ten thousand inscriptions which appear to contain mainly proper names, often borrowed from Latin or Greek; Etruscan is more often the recipient than the donor, though Etruscan also functioned as donor (e.g. historio in Latin from ister actor in Etruscan).

Some Ancients regarded Etruscan to be like no other language. Herodotus suggested Lydian affinities. Other affinities of Etruscan have been sought in Caucasian languages. If the Etruscans were not aboriginal to Italy, they migrated there from Asia Minor via Greece (perhaps as the Pelasgians).

Affinities have also been sought with other extinct languages, as Lemnian.

2.4. Lemnian is the language of a 6th century B.C. inscription of ten lines, discovered on the island of Lemnos.

2.5. Linear A is the exclusively Cretan predecessor of Mycenean Linear B. Linear A is still undeciphered (Middle Minoan, after 2,000 B.C.); it was not spoken in Greece as was Linear B, but only in Crete. Half of the Linear A signs correspond to those of Linear B, which was spoken and written after Linear A (Late Minoan). Though Linear B is now known to be an Indo-European language (Greek), it is expected that--once deciphered--Linear A will turn out to be unrelated to Indo-European languages.

2.6. Eteocretan is the name given to an undeciphered script of Crete. Eteocretan was supposed to be akin to Lycian by Herodotus.

2.7. Eteocyprian is also an undeciphered script, found in Cyprus.

2.8. Sumerian is the world's oldest known written language, with inscriptions dating from 3,100 B.C. Sumerian flourished in southern Mesopotamia during the third millenium B.C. until it was replaced (at least as a spoken language) by Akkadian, about 2,000 B.C. In written form, however, Sumerian continued to be used almost to the beginning of the Christian era. The impact of Sumerian on the development of the ancient Mesopotamian civilizations was entirely out of proportion to its restricted geographical distribution and small number of speakers.

"Unlike Assyria, Babylonia and Egypt, the histories and traditions of which are amply documented in the biblical and classical sources, there was nothing

to be found in non-Mesopotamian sources anywhere to make one even suspect the existence of the Sumerians in antiquity, let alone to appreciate fully the important role they played in the history of early civilizations.

"When the decipherment of cuneiform writing was achieved in the early decades of the 19th century, three languages written in cuneiform were discovered...Only after the texts written in Babylonian had become better understood did scholars become aware of the existence of texts written in a language different from Babylonian. When the new language was discovered it was variously designated as Scythian or even Akkadian, that is, by the very name now given to the Semitic language spoken in Babylonia and Assyria. As knowledge of the new language grew, it was given the correct name of Sumerian" (Ignace Jay Gelb, Encyclopaedia Britannica 21, 1964).

Of the several Sumerian dialects known, the most important are:

eme-KU (the official dialect)

eme-SAL (the dialect often used in hymns and incantations).

The Sumerian phonemic system (after Gelb) is as follows:

p	t			k		i	u
b	d			g		e	a
	s	ś	š		h		
	z						
m	n						
	l						
	r						

Sumerian abounds in affixation. Verbs and nouns are not distinguished qua major morphemes (i.e. roots); for example, /dug/ means speak as well as speech, the difference depending on affixation and syntactic environment.

2.9. Elamite (Anzanite, Susian, Hozi) is known from inscriptions dating from about 2500 B.C. This language, with no known ancient relatives or modern descendants, was spoken in an area corresponding to the modern Luristan and Khuzistan in Iran. Elamite was still spoken to some extent in the first century A.D. The earliest Elamite inscriptions are written in a unique script, often accompanied by an Akkadian version. A second group of inscriptions (16th to 8th century B.C.) is written in a cuneiform alphabet adapted from Old Babylonian; and a third, with Old Persian and Akkadian translations, dates from the 4th or 5th centuries B.C.

2.10. Speakers of Hurrian (Khurrian) inhabited an area called Mitanni (with Nuzi as its capital) which extended from northern Mesopotamia into Anatolia. Hurrian was written in Akkadian cuneiform. Tusratta (Dusratta) wrote a letter, about 1410-1375 B.C., to an Egyptian Pharaoh (Amenophis III) concerning the latter's marriage to the daughter of the Hurrian writer--the longest Hurrian text extant.

2.11. Urartaean (Vannic Haldean, Kaldic), spoken in the Mount Ararat area near Lake Van, is related to Hurrian. Information on this language is derived from two hundred inscriptions, written in Akkadian cuneiform dating from the ninth century B.C.

2.12. Hattic (Kattish) was spoken in Asia Minor, without known linguistic

relatives.

2.13. Kassite (Cossaean) was spoken in the Zagros Mountain area; no known relatives have been found for Kassite. This language is known from a bilingual glossary, after the 17th century B.C. and before the conquest of the Kassites by Alexander the Great.

2.14. Other languages formerly spoken in Asia Minor, without definitely known affinities, but attested by some inscriptions, include:

Sidetic

Carian

Tyrrhenian (possibly related to Etruscan)

Mysian (said by the Ancients to be a mixture of Lydian and Phrygian).

2.15. Still other non-Indo-European languages associated with the same general area are known chiefly from Greek traditions (but these are not to be confused with extinct Indo-European languages of Anatolia and the Mediterranean that are listed under 1, in Languages of the World: Indo-European Fascicle One):

Bithynian

Cappadocian

Cataonian

Cilician

Gergito-Solymean

Isaurian

Lyconian

Myriandynian

Paphlagonian

Pisidian

Pontic.

PIDGIN-CREOLES

3.0. If all languages are mixed, as has been said, this is tantamount to saying that each language is now in contact or has been in contact with at least one other language from whict it has borrowed. The traditional assumption is that languages in contact do borrow from each other, without exception; and that being in contact with one another is the natural condition of human languages, without exception. It is quite possible to imagine, however, that speakers of a single language, say Arctic Eskimo, may move into an uninhabited island, say Greenland, which then includes no other language. As long as the speakers of Arctic Eskimo dialects are in contact with each other while living in Greenland, they may borrow from each other--but then the result of this is surely not a mixed language, since there is only one in-coming language to begin with (Arctic Eskimo); rather, it is a kind of mutual adaptation known as dialect leveling.

The consequence of dialect leveling is uniform change over many settlements, even though each particular change is initiated in one settlement. But one settlement may move out of contact--as the Polar Eskimo did when they moved so far up the west coast of Greenland that they ceased having visits with West Greenlandic communities for centuries (and also lost the art of

making kayaks which they took north with them, as is known archaeologically). During these centuries, non-uniform change would go on: one set of changes in Polar Eskimo, and another set in West Greenlandic communities. These independent changes were interrupted, for European boats sailing in the Arctic in the Age of Discovery brought the Polar Eskimo and the West Greenlandic speakers together, thereby permitting dialect leveling to resume.

Contact is therefore the permissive factor, but attempts to communicate with one another is the decisive factor in dialect leveling.

Even across language barriers, willingness to learn the language of neighbors may result in language leveling, at least if the languages are closely related. Thus, after French and Provençal soldiers communicated in one army in two world wars, and after a whole generation of exposure to education in Standard French, speakers of Provençal have borrowed so heavily from French that Provençal may now be counted as a dialect of French. It is not that some Provençal speakers have taken the trouble to learn French, while others have not; nor is this an instance of replacive bilingualism; it is rather that the Provençal language itself has changed from being a separate language (whose speakers could not understand French) to becoming a dialect--albeit a divergent dialect--of French. It is Provençal that has changed in this contact which was one of bilateral bilingualism in two world wars (some French soldiers learning Provençal while, at the same time, some Provençal soldiers learned French), and continued after the wars as unilateral bilingualism (French taught in Provençal schools, but

Provençal not taught in French Schools).

Unilateral bilingualism is characteristically found the world over when languages are in contact, and such bilingualism characteristically fails to result in leveled languages. Bilateral bilingualism is rare; if it were more common, there might well be fewer language barriers in the world today and, by the same token, more instances of leveled languages.

Contact commonly permits leveling among dialects of the same language, as mentioned; less commonly, it permits leveling between two separate languages--at least if they are closely related--as also mentioned. But we are concerned with these two kinds of leveling as prolegomena to examining yet another kind--the only genuine kind of language mixing--not to be confused with the first two in which a given language changes but does not really mix in any genetic sense, since the changed language can still be traced as having descended from a single proto-language

(1) after contact has permitted leveling within a given language (dialect leveling);

(2) after contact has permitted leveling across a language barrier (when the two once-separate languages are closely related members of the same language family);

(3) however, after contact permits the immediate creation of a new language barrier in consequence of at least two (often more than two) languages mixing, then the new separate language so created cannot be traced as having descended from a single proto-language. This third kind of contact is well attested

since seafaring days in the Age of Discovery, and was perhaps formerly frequent in Mediterranean Europe and in the Most Ancient East.

When many people speaking language A come in contact with many people speaking language B, they will, under the appropriate cultural conditions, mix A and B to create language C, which is called a pidgin language. Pidginization is characteristically ephemeral. In the first generation, the speakers of the new language (C) are necessarily bilingual, since each C-speaker is a native speaker either of A (though some A-speakers do not learn C), or of B (though some B-speakers do not learn C). The children of the relatively few C-speakers are not necessarily bilingual in A or B; when the new language (C) is no longer restricted to the practical domains of their parents' use of it, and is spoken as the native language of monolingual children in all cultural domains, it is called a creole language instead of a pidgin language. The new language (C) is called a pidgin when all C-speakers are bilingual in it and in the native language of their forebears (A or B). This new language (C) may be called a creole instead of a pidgin in the second or third generation of its existence--as soon as some children of a new generation learn it as their mother-tongue (that is, become monolingual C-speakers). Or the new language (C) may be called a pidgin-creole in the general case in which some C-speakers have also a knowledge of languages A or B (and perhaps of several additional languages in multilingual cultures of New Guinea or acculturated Australia), though some C-speakers may be monolingual in C, in the same culture.

Instances of mixed languages of the pidgin-creole type are sometimes

assumed by the Ancients, who cite some of the languages listed above (2), though without evidence. It is possible that the then new juxtaposition or coexistence of written languages besides the continuation of of preliterate languages may have been stimulating for the more than occasional creation of a new language (C)beside the continuation of languages of the A and B type; at least it can be said that it is not at all possible to classify some of the formerly spoken languages of the ancient world according to language family affiliation (2, above), but it is possible to classify other formerly spoken languages of the same period in the same areas according to known language family affiliations, (as in 1, in Languages of the World: Indo-European Fascicle One).

However, the only sure evidence that one is dealing with a pidgin-creole language of the C type in the modern world is found when a known European language of the A or B type is the predominant donor of the vocabulary used by C-speakers who contribute features of structure from other languages which can rarely be identified. These unidentified structural influences are sometimes called the substratum. But this contributes a mystical label rather than additional information. Additional information about the non-Indo-European donor languages is sometimes found in later vocabulary accretions which are, however, often highly localized. Neo-Melanesian, for example, includes such vocabulary accretions in one part of New Guinea that are not used in another part of New Guinea (reflecting only the local languages of each part).

A remarkable feature of all modern European-based pidgin-creolesof the

C-type is the location of their speakers: all are spoken on or near seaboards. Another general feature is the plasticity of languages of the C type in contrast to what Sapir called the massive resistence to change in language (having in mind languages of the A and B type). Where C-speakers have enhancing contacts with English, Dutch, Portuguese, or German speakers, pidgin-creoles are reshaped so massively that the language barrier (between C and A or B) may be leveled. Such leveling is now complete in Hawaii, where many pidgin speakers continue speaking a kind of 'da kine talk' which is so reshaped under recent English influence that it is quite intelligible to monolingual English speakers. Much the same process is almost complete in Jamaica where Jamaican Creole--except for Bongo Talk of the elderly--is reshaped in more than one stylistic variety of English--from Standard to Quashie (substandard).

The language barrier between a pidgin-creole (C) and its predominant donor (A) may, on the other hand, be valued and consequently preserved by C-speakers who also know how to use language A (say Standard English), and use the Standard (A) or the substandard or really different pidgin-creole (C) in complementary situations called diglossia by Charles A. Ferguson (Word 15.2, 1959). In the general or typical instance of diglossia, language A (the Standard) would be used invariably for formal public occasions--e.g. official government activity, legal processes, academic ceremonies; for informal public occasions--e.g. conversation at concerts, market place--the Creole would be spoken. Formal but private occasions (receptions, introducing strangers) would call for the Standard, even though Creole is invariably used

in private informal conversations (between friends, lovers and kinsfolk).

In some areas where Creole languages are spoken--as in Surinam, parts of the Antilles, and in Louisiana--the official language is not the main donor. Surinam, for example, has an English-based Creole language; but Dutch is the official language of Surinam. Existing political (and official language) affiliations have here come after the creation of Creole languages. Occasionally, in these circumstances, it is not easy to find the European base for certain Creoles. One such Creole is Papiamento--called here a general Indo-European based Creole. Papiamento is spoken by 200,000 people in the Dutch West Indies (the islands of Curacao, Aruba, Bonaire) and probably also in the former Dutch possession, the Virgin Islands (just east of Puerto Rico), now administered by the United States and Great Britain. Albert Valdman considers Papiamento a Portuguese based Creole; according to Navarro Tomás and Van Wijk (R. W. Thompson in Creole Language Studies II, 1961), Papiamento is a development of the West African slavers' jargon (i.e. Afro-Portuguese); Robert Hall and William A. Stewart speak of Papiamento as an outgrowth of pidgin Spanish and Portuguese with heavy borrowings from Dutch.

In the discussion of languages of the C type (pidgin-creoles), we distinguish below between those that are Portuguese-based (3.1), those that are French-based (3.2), and those that are English-based (3.3), with Neo-Melanesian (3.4) taken as a special instance of the latter, for it is particularly diversified, but still not as diversified as are Surinam Creoles (3.5).

Additional pidgin-creoles, most of which are extinct, are then listed without discussion (3.6).

3.1. Portuguese-based pidgin-creoles are spoken in areas extending from the Cape Verde Islands off the coast of Africa, to mainland parts of Africa, to the Indian sub-continent, to Southeast Asia, Macao, the China coast (Hongkong, Shanghai), and to the Philippines, and perhaps as far in Polynesia as Hawaii. It is likely that an Afro-Portuguese pidgin-creole diffused with Portuguese territorial expansion around the Afro-Asian coastline from Cape Verde to Japan. The same Afro-Portuguese pidgin-creole may have diffused to the Carribbean and French possessions in the Indian Ocean.

African Afro-Portuguese:

Cape Verde

São Tomé

Portuguese Guinea

Senegal

Asian Afro-Portuguese:

Indo-Portuguese

Molucca

Tugu (Java)

Macanese

Caviteño

Ermitaño

Zemboangueño (Chabacano) with a subdialect development know as Davenefio.

When first discovered in 1460, the Cape Verde Islands were uninhabited. The first settlers were given monopoly of the slave trade of the Guinea Coast (with the Cape Verde archipelago serving as the center of assembly and export of slaves); this slave labor attracted large numbers of Portuguese settlers. Chronic overpopulation caused constant re-migration of many of Cape Verde's Portuguese, mulattoes, and freed slaves to other Portuguese possessions. Such re-migration was accelerated at the termination of the African slave trade (1876); many Afro-Portuguese speakers went as far afield as the United States (Rhode Island, Massachusetts) and across the Pacific to Hawaii.

The population of the Cape Verde archipelago has been decreasing rather than increasing; it is now 150,000 (1960). Each of the ten main islands of Cape Verde archipelago has its own dialect of Cape Verde Creole; there is little contact between islands.

Afro-Portuguese has spread with little change to other sections of Portugal's African territories--São Tomé (8,000 speakers), Senegal, and Portuguese Guinea (50,000 speakers); in some instances in these areas, the pidgin-creole was spread by runaway slaves.

Afro-Portuguese influences reached into South Asia and Southeast Asia as colonists from Portuguese territories of Africa re-migrated to Goa, Java, and the Moluccas in the 1500's. Asian-Afro-Portuguese pidgins developed, which were replaced by Portuguese and Indian dialects in India; these pidgin-creoles continued in parts of Macao, the Moluccas and Java. In the city of Macao, the Macanese Creole coexists with Portuguese; Macanese is structural-

ly similar to Portuguese-based pidgin-creoles in other parts of Africa and Asia.

For Macanese spoken in Hongkong, Thompson (op. cit.) says:

"This is the home and community language of up to 4,000 people of mixed ethnic descent, whose ancestors came from Macao after the Island of Hongkong was occupied by the British in 1841. It is the best preserved of all the varities of Macanese on the China Coast; that of the city of Macao itself has grown progressively more like metropolitan Portuguese during this century. The speakers of the Hongkong dialect are generally tri-lingual, speaking Macanese, Cantonese and English, in which last languages they receive their education. There is little evidence of Cantonese influence in the structure and lexicon of their mother tongue which seems to have come to China ready-made. Its structural similarities to the Malayo-Portuguese dialects of Malacca and Java, to the Indo-Portuguese complex and the Portuguese Creoles of West Africa (even including the language used by Gil Vincente's farcical Africans) are much more numerous than the occasional resemblances to Cantonese structure.

"The verb in Macanese generally consists of one single word-base or stem, which alone or in conjunction with a particle can express mode, aspect or tense. The stem alone expresses habitual aspect or past tense as in Jamaican or Haitian Creole: éle falá makísta he speaks, spoke Macanese. The durative or non-completive aspect of the verb is expressed by the particle ta followed by the stem: sol ta subi, the sun is rising; yo ta skrevé kwando ele ven, I was writing when he came; ele ta čurá she is weeping; nos ta ven

kwando veŋ čuwa we were coming when the rain started.

"The punctual or completive aspect (past tense) of the Macanese verb makes use of a perfective particle ǰa: ele ǰa veŋ he came; osé ǰa olá ele ǰuntado ko nos you saw him with us; nos ǰa vai makao we went to Macao; yo ǰa peská uŋa grande peše I caught a large fish.

"The contingent or future mode is formed by the particle logo and stem: ele logo kazá ko María he will marry Mary; nos logo vai merkado we shall go to market; yo logo falá I shall speak; dos muler logo vai ǰuntado ko elotro-sa kyansa the two women will go with their children; ilotro kerá olá kwanduŋa logo pode fazé ome irá kapote pemero they wanted to see which one would be able to make the man take off his cloak first. In the negative a special particle nadi(<Port. não ha de) is used instead of logo. It also occurs in the dialect of Malacca and in the Indo-Portuguese dialects."

Bloomfield mentioned that a Spanish jargon was formerly spoken in the Philippines (Language, 1933). Keith Whinnom found three Creole languages spoken in the Philippine Islands in 1942, which he names after the principal localities where they were spoken, by 18,000 and 12,000 and 1,300 speakers respectively:

Caviteño

Ermitaño

Zamboangueño (Chabacano).

A fourth Creole mentioned is Davaneño (3,000 speakers) which is a dialect of Zamboangueño. Whinnom is of the opinion that the Creoles were based on an

Indo-Portuguese pidgin spoken since 1500 A. D. in the Portuguese possessions of the East Indies and India. A Spanish garrison from this zone had been removed to the Philippines in 1658.

Taylor (Word 13 (1957), p. 493), after Whinnom, gives the following tentative phonemization of all four Portuguese-based Creoles of the Philippines:

p	t	č	k		
b	d		g		
	s	š	h	i	u
m	n	ñ		e	o
	l	ɬ		a	
	r			/ ˈ/ stress	
w	y				

3.2. French-based pidgin-creoles are spoken in the following areas; numbers of speakers are given in parentheses:

Circum-Carribbean area:

Louisiana (80, 000)

Haiti (4, 345, 948)

French Guiana (35, 000)

Lesser-Antilles (800, 000)

Guadeloupe

Les Saintes

Martinique

Dominica

Trinidad

Grenada

St. Lucia

Afro-Asian area:

Réunion (250,000)

Madagascar

Mauritius

Rodriguez (17,000)

Chagos

Agalega

Seychelles (42,000)

South Vietnam (Annamito French)

French-based pidgin-creoles of the Caribbean are said to be mutually intelligible. According to Stewart, "speakers from opposite ends of the Caribbean--from Louisiana in the north and French Guiana in the south--are able to converse together with a minimum of misunderstanding. Similar cases, though on a much smaller scale, are common in other Creoles as well." (Study of the Role of Second Languages, Washington (1962), p. 65).

Louisiana French Creole is represented by that of St. Martin Parish (Raleigh Morgan Jr., Structural Sketch of St. Martin Creole, AL 1:8, 20-4, 1959), the phonemes of which are:

p	t	č	k				
b	d	ǰ	g			i	u
f	s	š		h		e	(ə) o
v	z	ž				ɛ	ɔ
m	n	ñ				a	ɑ
	l	r				ɛ̃	ɔ̃
w	y						ɑ̃

Today Haiti is one of the most densely populated of the American Republics (population 4,345,948 in 1962). French is the official language but 90 percent of Haitians speak Creole. Albert Valdman estimates that almost the entire phonology and lexicon of Haitian Creole is Romance; the structure is not. Whole sentence types occur which do not occur in French.

The following phonology for Haitian Creole is Albert Valdman's (From Creole to French in Haiti):

p	t	č	k			i	u
b	g	ǧ	g			é	ó
						è	ò
f	s	š		h			a
v	z	ž	r			ẽ	õ
m	n	ñ					ã
	l						
w	y						

This scarcely differs from that of the McConnell-Laubach-Hall phonemicization (Robert A. Hall Jr., Haitian Creole: Grammar, Texts, Vocabulary,

AA Memoir 74, 1954).

The Creole spoken in Dominica (Lesser Antilles) makes much the same phonemic contrast made in Haitian Creole. For example there are three linear distinctions among nasals in both (but palatal /ñ/ in Haitian and velar /ŋ / in Dominican). Both show two coexistent vowel systems, one for oral vowels, one for nasalized vowels. These vowel systems are quite symmetrical in Haitian (see above); they are wholly symmetrical in Dominican whose oral vowels contrast front, central (or unrounded) and back (or rounded) at two tongue-heights, and whose nasalized vowels contrast front-back at two tongue-heights, according to Douglas Taylor (Word 3.173-9, 1947; 7.43-59, 1951).

French-based creoles are spoken on islands facing Africa in the Indian Ocean and in Southeast Asia and in Melanesia (New Caledonia). Just as the French Creoles of the Caribbean--from Louisiana to French Guiana--are mutually intelligible dialects, so too are most of the French Creoles from East Africa to Vietnam and perhaps even to New Caledonia.

Réunion is the first island of the Mascarene archipelago, five hundred miles west of Madagascar. The population exceeds 300,000 of whom the majority speak an Austronesian-Afro-French Creole. The extra-French influence in this creole is from Malagasy and Bantu. When France took possession of the island in 1638, it was uninhabited; the inhabitants of Réunion are descendants of migrants. Between 1634-1848 the French introduced African slaves to work their sugar plantations. After 1838, Indians and Chinese (40,000 today) were brought in as indentured laborers.

It is certain that Réunion Creole is spoken on Madagascar (where the Austronesian language, Malagasy, is the lingua franca, and the official language) since the population explosion in Mauritius and Réunion is causing constant migrations to the relatively under-populated Madagascar.

The first French settlers in other Indian Ocean islands facing Africa (Mauritius, Rodriquez, and Seychelles), in Southeast Asia (Vietnam), and in Melanesia (New Caledonia) all came from Réunion in the 18th and 19th centuries. They may well have brought with them some variety of Réunion Creole.

3.3. The following English-based Creoles are discussed here:

Bongo talk (Jamacia)

Gullah (coastal and offshore island facing South Carolina and Georgia)

Krio (Sierra Leone)

Surinam Creole (Saramaccan, Djuka, Sranan, Aucaan, and Matuwari) are discussed below (3.5), as is Neo-Melanesian (3.4).

In Jamaica, Bongo talk continues despite the fact that official and educational policy has been to foster Standard English and to refuse to accept the presence of pidgin-creole even when reported, as it is in Creole Language Studies II (1961) pp. 83-4. Although the lexicon is predominately English, a

high proportion of African loan words, and phonological and grammatical features which are non-English show that Bongo talk includes African admixture. It is unintelligible to non-speakers (i.e. to other Jamaicans); a few elderly people in isolated parts of Jamaica speak Bongo Talk as their only language.

Gullah is the Creole spoken on the islands off South Carolina and Georgia by descendants of former slaves (also called Gullahs or Geechees). Its vocabulary has been traced to several West African languages (see L. D. Turner, Am. Dialect Soc. (May 1945 and April 1948)--Mende, Temne, Twi, Hausa, Yoruba, Ibo, etc.--which were the languages spoken by African slaves imported in the 17th, 18th, and 19th centuries. The influence of West African languages and dialects is also obvious in Gullah's phonology and morpho-syntax. The following phonemicization--a tentative one--is after Turner (1948):

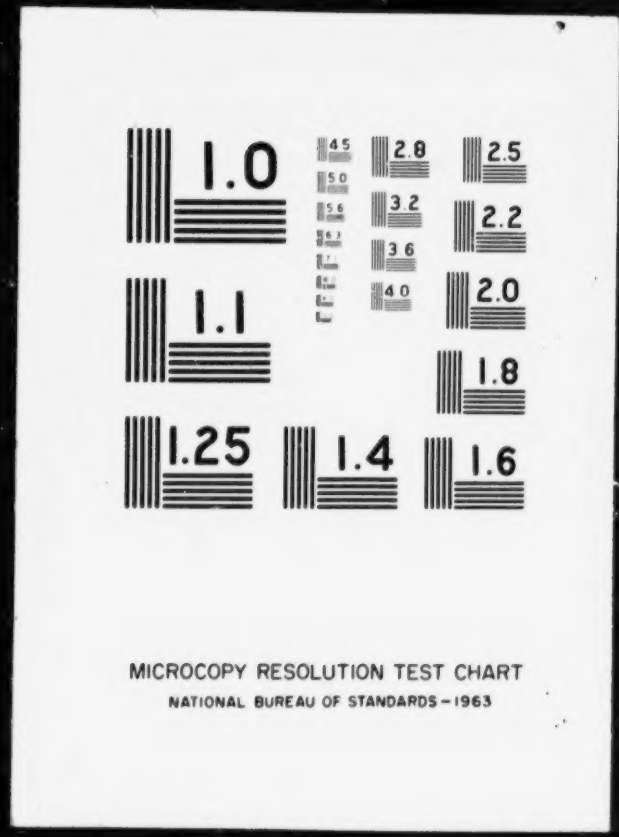
p	t	c	k	kp			
b	d	ʒ	g	gb		i ɨ	u
f	s	š		h		e ə	o
						ɛ ą	ɔ
v	z					ĩ	ũ
m	n	ñ				ẽ	
	l r					ã	
w	y						

Plus prenasalization for all stops

Plus glottalization for /p t k š. . ./.

Data on Sierra Leone Creole is based on Berry (English Loanwords and Adaptations in Sierra Leone Krio, in Creole Language Studies II, 1961).

2 OF 2
ED
010355



Krio is the mother tongue of 20 or 30 thousand people in the port city of Freetown, and of the satellite villages in the African hinterland of Freetown. It is mainly English-based today; but includes many admixtures, especially of pidgin of the 19th century spoken by returning freed slaves--many from Jamaica (hence Krio may be related to Jamaica's Bongo Talk), by indigenous Africans, by Europeans, and by settlers--'Black Poor' from Nova Scotia and Britain and Maroons, offspring of Portuguese-African and British African unions. Intensive 19th century borrowing from African languages--notably from Yoruba, Mende and Temne--reshaped Krio just as 20th century borrowing from English is again reshaping it.

The language situation in the Freetown area is a multilingual one. The typical Freetown urbanite has Krio as his primary language and English as his secondary language. Aboriginal African languages are spoken in the interior, away from the Freetown seaport area, and also by migrants to Freetown from the interior.

The phonemes of Freetown Krio appear below (with phonemes used only in African loans enclosed in square brackets):

p	t	č	k	[kp]	i	u	[i̩ u̩]
f	s	š			e	o	
m	n	[ɲ]	ŋ	[ŋm]	ɛ a ɔ		[ɛ̩ a̩ ɔ̩]
	l	r					
w	y						

Plus voicing for stops and fricatives.	Plus two degrees of stress Plus / ˋ ˊ ˇ ˆ / tones (English stresses are phonemicized as high and low tones).

Voiceless stops are strongly aspirated, with frication occurring at point of articulation. [kp] and its voiced counterpart [gb] are coarticulated stops which have been adopted from Yoruba, Mende and Temne . Before / a/ , the velar stops / k/ and / g/ are usually palatalized and this is phonemically represented by a vowel cluster: cat /kiat/ girl /gial/ .

Nasals have complex variants: after a vowel, / m, n, ŋ/ vary freely with nasalization of the vowel (e.g. / s ɛ n s/ / s ɛ̃ s/ sense); these three nasals are also neutralized in preconsonantal position, occurring either homorganic with the following consonant or as nasalization of the preceding vowel.

3.4. Neo-Melanesian, or Pidgin English, is spoken in the Australian Territory of New Guinea (including the Bismark Archipelago), in the Solomon Islands and adjacent islands. Robert A. Hall, Jr. (in the introduction to F. Mihalic's Grammar and Dictionary of Neo-Melanesian, 1962) estimates the total Neo-Melanesian speech community to be nearly a million. Neo-Melanesian is restricted to the coastal regions; only about 5 percent of the males in the New Guinea highlands know Pidgin.

The vocabulary is 88 percent English, including English from sub-standard dialects, and including compounds not made in standard English. The remaining 12 percent of the vocabulary is from Melanesian, Malay, Polynesian, German, Chinese, and Spanish.

The Neo-Melanesian phonology which follows is after Mihalic:

p	t	k		i	u
(b)	(d)	(ǰ) (g)		e	u
					ɵ
(f)	s		h	ɛ	o
(v)				æ	a ɔ
m	n	ŋ			
	(l)	r			
w	y				

Consonants enclosed in parentheses show restricted distribution; they do not occur in some dialects. Voiced stops never occur finally; thus the feature of voiced/voiceless is neutralized in this position. However, in some Neo-Melanesian dialects, the phonemic system of the speaker's native language allows no final consonants; where consonants occur word-finally, a vowel or a syllable is added and the final /b, d, g/, now prevocalic, contrast with /p t k/.

The phoneme /v/ is rare; it becomes /p/ or /b/ in many dialects.

The phoneme /f/ is rare; /p/ is characteristically substituted for English /f/ in Neo-Melanesian.

The phoneme /ǰ/ occurs rarely, and then only initially. The palatal affricate varies between affricate and fricative, as /ǰas/ or /zas/ for judge. Voiced stops are frequently pre-nasalized, especially by Melanesian speakers. Consonant clusters are limited in many areas to /tr, br, dr, gr, kw, tw, sw, nk/ and may otherwise be simplified to single consonants. Many English consonant clusters such as /st, sk, sp/ are separated by a vowel.

Examples of contrasts between consonants now follow:

/p/ contrasts with /b/:

bik big

pik pig

/p-b/ contrast never occurs finally:

clap club

slip sleep

/t/ contrasts with /d/:

dok dog

tok talk

/t-d/ contrast never occurs finally:

bet bed

blut blood

het head

hot hot

kot court

/k/ contrasts with /g/:

kos course

gam spirit

kam come

gam cowrie shell

/k-g/ contrast never occurs finally:

bek bag

dok dog

pik pig

sik sick

/l/ contrasts with /r/ in some dialects:

long at, by, to

rong wrong, fault

/l/ and /r/ are both sometimes flapped (and the distinction is lost in some dialects):

lait	light	lat	lathe
rait	write	rat	rat

/f/ is listed as an initial phoneme by Mihalic for 22 words--e.g. famili family, fut foot, faktori factory.

/ǰ/ is a rare phoneme; the most common word with /ǰ/ is ǰip jeep, car.

/w/ and /y/ are consonants by virtue of their distribution before vowels as in kawawar ginger root, wan one, wara water, yang young, yu you, yia year.

Neo-Melansian is spoken as a second language by native speakers of several hundred Papuan and Melanesian languages; often the phonology of a given native language is reflected in the spoken Pidgin. However, Hall and Mihelic set forth certain adaptations which occur more or less universally when a speaker of a Papuan or a Melanesian language speaks Pidgin-English:

(1) /b, d, g/ are prenasalized in initial and intervocalic position as in mbuk book, sindaun sit down, nongut no good.

(2) In final position, when word final sequence is nasal + stop, the stop is lost, as in lam lamp, pain find;

(3) A sibilant and stop or nasal in word initial position are separated by / i/ (CVC), as in sitik stick, sikin skin, sinek snake.

There are sixteen highly recurrent minor morphemes in Neo-Melanesian; they are listed here without co-occurrence restrictions being cited--the most often used function is given below in the gloss.

-im	transitive
i-~ in-	predicate marker
bilɔŋ	possessive, connective, relation marker
no	negative marker
inap	ability
ken	ability, future
lɔŋ	preposition (by, to, at, because, connected with, etc.)
pinis	completive
tasol	intensive
tru	superlative
hap	partitive
nau	and
mas	must
na	connector
no	or
laik	future

3.5. In Dutch Guiana or Surinam, Dutch is the official language, and is spoken beside Chinese, Hindi-Urdu, Javanese, several American Indian languages and several Creoles. In the 1650's, the Surinam section of Guiana was settled by the English, who were joined by some Portuguese who were evicted from Brazil in 1664 In 1667, the Dutch invaded the colony, and have not since relinquished their governing power. Previous to 1863, European settlers imported African slaves to work on their plantations. Many of these enslaved workers escaped into the jungle interior (hence 'bush negroes'), where they live in partial isolation today, with their lands and power of restricted self-government secure under a treaty. After emancipation, the coastal negroes migrated to urban areas, or became independent peasantry; and Chinese, Hindu-Urdu, American Indians, and Indonesians were introduced as indentured plantation workers. Today, the descendants of Africans are in the majority.

It is certain that the slaves introduced to Surinam were from West Africa, but there is no evidence to indicate whether they spoke a common language, as Afro-Portuguese, or a number of distinct African languages. Probably the latter, since dominant coastal people speaking Afro-Portuguese raided the interior for slaves; Afro-Portuguese was not spoken in the areas in which the slave raids took place. There is still disagreement as to whether pidgin-creoles of Negro-English, Negro-Portuguese, and Negro-Dutch existed independently in Surinam in the past and, after some time, merged to form the existing languages.

Jan Voorhoeve (Word 15. 3:436 (1959); Creole Language Studies II, (1961), 99-106) gives the following account of Creole languages and dialects of Surinam. Creole became an important language among African descendants; it also became the language of many of the newly arrived Asian migrants and their descendants. The combined reports of Jan Voorhoeve and Douglas Taylor (Word 15. 3:487, 1959), and J. J. M. Echteld (The English Words in Sranan, Groningen, 1962) make possible the following list of Surinam pidgin-creoles:

Saramaccan (Jew Tongo)

Upper Saramaccan

Lower Saramaccan

Djuka

Sranan (Negerengels, Nenger-Engels, Taki-Taki, Kriono, Nengre-Tongo, Tongo)

East Sranan

West Sranan

Aucaan

Matuwari.

Dialects of the Saramaccan language are spoken along the Surinam and Saramaccan Rivers. Djuka dialects are spoken along the Marawijne River, and Sranan dialects are spoken on the coastal belt.

Major dialect differences exist between speakers of the eastern part of coastal Surinam--the East Sranan dialect (an old plantation area)--and the western part of coastal Surinam where the West Sranan dialect is spoken

(a new plantation area). In the latter, English loanwords are plentiful; allophones of phonemes as well as sandhi-phenomena are strikingly different from those of the East Sranan dialect. It may well be that the differences are the result of historical accident. When England stopped the African slave trade, Surinam was just opening up the new plantation area. Slaves that were outlawed in English possessions were shipped to Surinam from the Caribbean Islands; these relocated slaves had already learned an English-based pidgin-creole.

Speakers of Sranan can learn to speak Djuka quite easily, because of vocabulary similarities between Sranan and Djuka. Sranan and Saramaccan, however, are quite divergent in form and content. The two differ in vocabulary (with Saramaccan having more Portuguese and African words); and in phonemic systems (with Saramaccan's being more African-like--Saramaccan is a tonal language, Sranan has stress; Saramaccan has seven vowels, Sranan has six; Saramaccan has coarticulated stops and a set of prenasalized stops, which are absent in Sranan; and also absent in Sranan is the f/v and s/z contrast found in Saramaccan). Saramaccan and Djuka are the languages of the 'bush negroes' mentioned above, while Sranan is that of the freed slaves and their descendants.

The following gives the phonemic contrasts made in Saramaccan Creole, according to Jan Voorhoeve (An Orthography for Saramaccan, Word 15.436-45, 1959):

p	t	t^y	k	kp		
b	d	d^y	g	gb		
f	s				h	
v	z					
m	n	n^y				
	l					
w	y					

i		u
e		o
	a	

Plus nasalization and length for all vowels;
Plus tone: / ˊ/ high
/ ˋ/ low
/ ˇ/ rising
/ ˆ / falling

Plus prenasalization for all voiced stops.

Vowel clusters, but not consonant clusters, occur. Tone sandhi occurs on word final vowels, which always have low tone. If a final vowel has high tone it is changed to low. In utterance final positions, all low tones preceding a high tone vowel are changed to high tone. Thus /àgènaˊ/ snake, becomes /àgèná/ in /mì sí wán àgèná/ I saw a snake; while /bàbúnù/ baboon becomes /bábúnù/ in /mì sì wán bábúnù/ I saw a baboon.

The following words exemplify the contrasts: pái father-in-law, bái to buy, mbata deer, agba chin, agbagba balance on head, akpo arrowpoint, ahalakpaka cockroach, a fíti hěn it fits him, aviti drill, vívo alive (of people), wéki to awaken, tii to steer, di the (sg.), ndika fish trap, asédu personal name, azéma vampire, foló flower, t^yumá to burn, d^yúnsu promptly, presently, nd^{y}u ground nut (species of), n^{y}ún^{y}u brand new, kó to come, gó to go, ngaku to stutter, hón to uproot, hái to pull, héi high, hulu to examine, wái to blow, wéi tired, wulu to wriggle, haun jealous, de they, dě to be, dá to give, ku with, déé the

plural, gĩĩ green, sĩtà to tie, sẽ́ shame, vĩ́tu wind, wã́ one, sṹ swim.

The following gives the phonemic contrasts made in Sranan, according to Robert A. Hall, Jr., (The Linguistic Structure of Taki-Taki, Language 24, 1948, pp. 92-116):

p	t	k				i	ü	u
b	d	g				e	a	o
f	s	x	h					
v	z							
m	n	ŋ						
	l	r						
w	y							

Plus stress; and nasalization for /i u e o a/

Consonant clusters of two members occur; all consonants may occur word initially (except /ŋ/) and finally (except /h/).

Sranan is an English-based Creole. The vocabulary consists mainly of words borrowed from West Africa, Portuguese, English and Dutch languages, with lesser accretions from Hebrew, French, Amerindian, Indian, Indonesian and Chinese. The following vocabulary strata are distinguishable:

(1) an African-English-Period in the 18th century which included some few Portuguese loans;

(2) an African-English-Dutch period in the 19th and 20th centuries, with Dutch loans gradually replacing previous loans; borrowing from Dutch was without assimilation to the Sranan phonemic system. The phonemic system of Sranan is Germanic-based, even to the extent of /h/ and /ŋ/ complementary

distribution, which is a typological featurecommon to all Germanic languages.

3.6. The following group of pidgin-creoles are listed with little comment. Most are extinct. A list of extinct pidgin-creoles attempting to be exhaustive would probably be much longer than the following:

Lingua Franca: the original lingua franca of Medieval times, with Provençal as the main European donor.

Beach la Mar: the 19th century pidgin of whalers, sealers, traders, and blackbirders of the Pacific ports, based mainly on English.

Russonorsk: spoken by Arctic Russians andScandinavians at the turn of this century; probably also spoken by Lapps.

Chinook Jargon: a North Pacific Coast Chinook-based pidgin, Chinook Jargon was developed by English and North American Indian languages flanking Chinook in the 17th century; for a time, it was used as a creole, as the native language of descendants of French voyageurs and Indian women.

Business English: the 19th century pidgin of Asian ports dominated by Chinese importers, exporters, businessmen, and coolies; English-based, with Chinese structure.

Sabir: the lingua franca of the ports of the Mediterranean, based on French, Spanish, Italian, Arabic, and Greek.

Also cited with little comment are Afrikaans and Fanagolo, both spoken in South Africa. Afrikaans is spoken by 2,250,000 white South Africans and about 750,000 Africans. It is sometimes supposed that Afrikaans developed from a creolised Afro-Dutch pidgin. The vocabulary is over 90 percent Dutch,

but the grammatical structure is Bantu-like in appearance, with simplifications resulting in the lack of grammatical gender among nouns, the use of one form of the verb for all persons with tenses formed by modals, the absence of strong and weak declensions characteristic of Germanic languages. Afrikaans is today mutally intelligible with Dutch, and hence classifiable as a modern Dutch dialect.

Fanagolo is the lingua franca used among relocated Africans speaking different languages in the mining areas of Johannesburg. The Johannesburg cultural situation favors the formation of a pidgin-creole; the status of Fanagolo, however, and its linguistic relationships with African languages, with English and with Afrikaans, is not known.

We list in other fascicles lingua francas (without raising the question as to whether a given lingua franca is a pidgin-creole or a language which can be shown to have descended from a single proto-language)--especially under 0 in Languages of the World: African Fascicle One.

There remain many instances in which secret languages or poorly learned dialects of languages are mistaken as pidgins. Thus, both Pachuco and Trader Navaho in the American Southwest have often been thought of as new jargons or pidgins. Trader Navaho is a kind of Navaho--poorly learned Navaho. And Pachuco may be thought of as a recently created Spanish dialect. Pachuco is a kind of Mexican Spanish that is not intelligible to monolingual speakers of Spanish. One view is that, beside the reduction of language barriers in the Southwest, Pachuco represents the creation of a new language barrier, a new

language. The intent of the young Mexican Americans who speak Pachuco appears to be to become unintelligible to their elders who, often enough, speak only Spanish, or good Spanish besides inadequate English. Since the innovators of Pachuco belong to the grandchild generation of three generation families, they are already bilingual (English-Spanish). Since they know English they are able to render into Spanish the literal or word-by-word translation of English idioms (calques). For the hearer who knows both English and Spanish, such calques would not take long to fathom. But the parents or grandparents of the Pachuco speakers know no more English than Anglo-Americans know Spanish. Hence the calque ingredient alone in Pachuco would serve the purpose of making this deviant Spanish a secret Spanish. Another main ingredient of Pachuco is slang--the use of known Spanish words with unknown or rare referents (compare English slang <u>to dig,</u> in the sense of comprehending the essential point). This kind of slang, as well as any reliance on introducing rare or archaic words from the same language will insure only a very temporary secrecy; we are all accustomed to inferring referents of new or unknown words from context. So we conclude that among the various devices for secrecy in Pachuco the use of calques is the most effective, because calques in Pachuco are difficult to infer from context by any except bilingual English-Spanish speakers.

The function of Pachuco speaking varies from Tucson to Phoenix and varies also between cities close to the Mexican border (from Texas to Los Angeles) and cities some distance removed from Mexico, as Flagstaff. The

Mexican border Pachuco is (or perhaps was, since its fashion may be waning) associated with daring or even delinquent youth who dress for the part (Pachuco dress) and self-consciously speak Pachuco as a symbol of their role in a design of living which fits neither Mexican American nor Anglo-American values, but may well be transitional between the two.

Pachuco can scarcely be classified as a separate language, even though a language barrier exists between it and monolingual Spanish speakers, and between it and monolingual English speakers; but bilingual English-Spanish speakers should be able to infer messages in Pachuco utterances from extensive contexts. Pachuco may be classified as a second dialect (analogous to 'second language') of certain bilingual Spanish-English speakers who wish to remain incommunicado with monolingual speakers in the same speech community.

To summarize, there are--besides secret languages and imperfectly learned languages (loosely spoken of as pidgins)--genuine pidgin-creoles or mixed languages. More is known about those which are based on Indo-European languages than those based on other languages. Finally, special mention should be made of two of the latter, Police Motu and Sango.

William J. Samarin has appreciated the sociolinguistic importance of Sango, spoken as a pidgin-creole by about a million people in the Central African Republic, whose official government-sponsored and missionary-sponsored language is French. But the non-elite who do not know French turn to the Creole Sango; and this Creole serves to conceal the rural ("bush people")

origins of its speakers. The consequence of this concealment is an extraordinarily diverse language. For example, the Sango spoken by a speaker whose native language is Banda avoids borrowing from Banda, while Sango speakers having other native languages may select the Banda borrowings. Hence, avoidance selection results in diverse usage out of the total lexical resources of the Creole Sango. Many languages are donors to the lexical resources of Creole Sango, which is the only Central African Republic language used in radio broadcasts, and is widely understood but spoken differently by the speakers of each different native language, for reasons mentioned above. The Creole Sango is based on Ngbandi which belongs to Group III of the Eastern section of the Adamawa-Eastern phylum.

Police Motu, based on Motu (an Austronesian language spoken in the central coastal area of the Central District, Papua), is a pidgin language with 'a relatively small working vocabulary'; the non-Motu words are mostly from English (S. A. Wurm and J. B. Harris, Police Motu, Linguistic Circle of Canberra Publications, Series B, No. 1, 1963). Police Motu is used as lingua franca over a much wider area of Papua than that in which Motu is spoken; native policemen in Papua are given training in Police Motu, and 'the bulk of Papuan [radio] broadcasting is done in it' (Capell, 1962).

The Following Abbreviations Will Be Used

AA	American Anthropologist
ACLS . . .	American Council of Learned Societies
AES-P . . .	American Ethnological Society, Publication
AL	Anthropological Linguistics
APS-P . . .	American Philosophical Society, Proceedings
APS-T . . .	American Philosophical Society, Transactions
BAE-B . . .	Bureau of American Ethnology, Bulletin
BAE-R . . .	Bureau of American Ethnology, Report
CU	Columbia University Contributions to Anthropology
IJAL	International Journal of American Linguistics
IUPAL . . .	Indiana University Publications in Anthropology and Linguistics
JAF	Journal of American Folklore
JSAP	Journal de la Société des Américanistes de Paris
Lg	Language
RCPAFL . .	Research Center Publications in Anthropology, Folklore and Linguistics
SJA	Southwestern Journal of Anthropology
SIL	Studies in Linguistics
TCLP . . .	Travaux du Cercle Linguistique de Prague
UMPL . . .	University of Michigan Publications, Linguistics
UCPAAE . .	University of California Publications in American Archaeology and Ethnology
UCPL . . .	University of California Publications in Linguistics
VFPA . . .	Viking Fund Publications in Anthropology
WDWLS . .	William Dwight Whitney Linguistic Series

CONTENTS

IBERO-CAUCASIAN AND PIDGIN-CREOLE FASCICLE ONE

0. Scope note 2

1.0. Caucasian languages 5

1.1. Northwest Caucasian 10

1.2. Northeast Caucasian 13

1.3. South Caucasian 21

1.4. Caucasian structures 23

2. Modern Basque and extinct languages of the Mediterranean and the Most Ancient East 31

3. Pidgin-Creoles 39